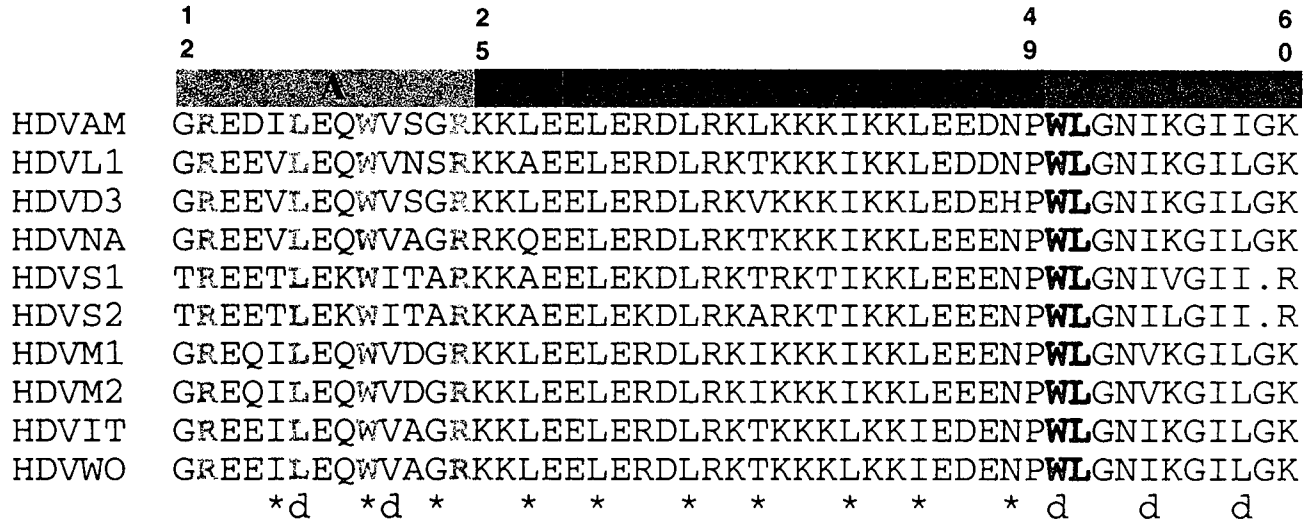
1
2
2
5
4
9
6
0
A
HDVAM GREDILEQWVSGRKKLEELERDLRKLKKKIKKLEEDNPWLGNIKGIIGK
HDVL1 GREEVLEQWVNSRKKAEELERDLRKTKKKIKKLEDDNPWLGNIKGILGK
HDVD3 GREEVLEQWVSGRKKLEELERDLRKVKKKIKKLEDEHPWLGNIKGILGK
HDVNA GREEVLEQWVAGRRKQEELERDLRKTKKKIKKLEEENPWLGNIKGILGK
HDVS1 TREETLEKWITARKKAEELEKDLRKTRKTIKKLEEENPWLGNIVGII.R
HDVS2 TREETLEKWITARKKAEELEKDLRKARKTIKKLEEENPWLGNILGII.R
HDVM1 GREQILEQWVDGRKKLEELERDLRKIKKKIKKLEEENPWLGNVKGILGK
HDVM2 GREQILEQWVDGRKKLEELERDLRKIKKKIKKLEEENPWLGNVKGILGK
HDVIT GREEILEQWVAGRKKLEELERDLRKTKKKLKKIEDENPWLGNIKGILGK
HDVWO GREEILEQWVAGRKKLEELERDLRKTKKKLKKIEDENPWLGNIKGILGK
*d *d * * * * * * * * d d d

Figure 1

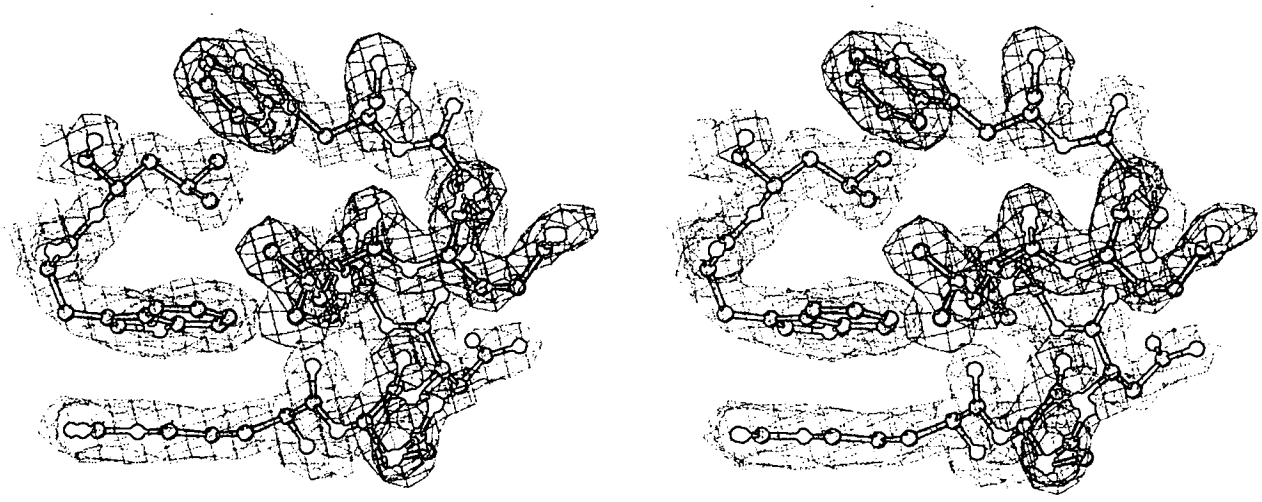

Figure 2

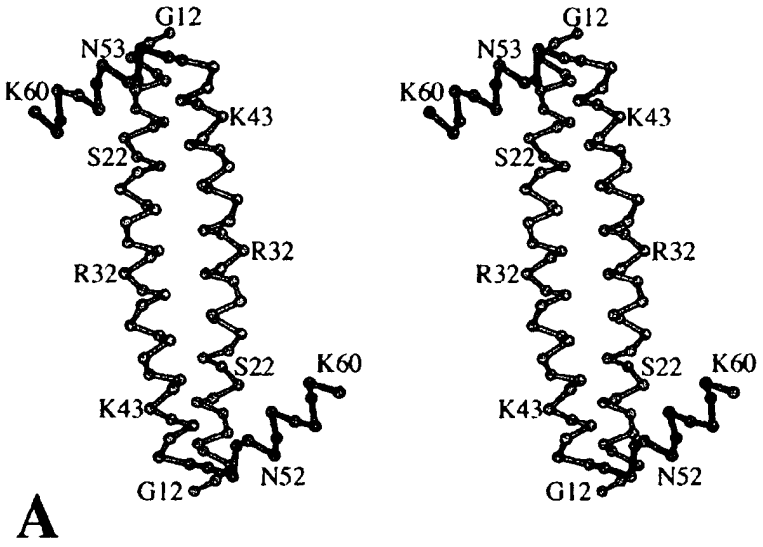

Figure 3A

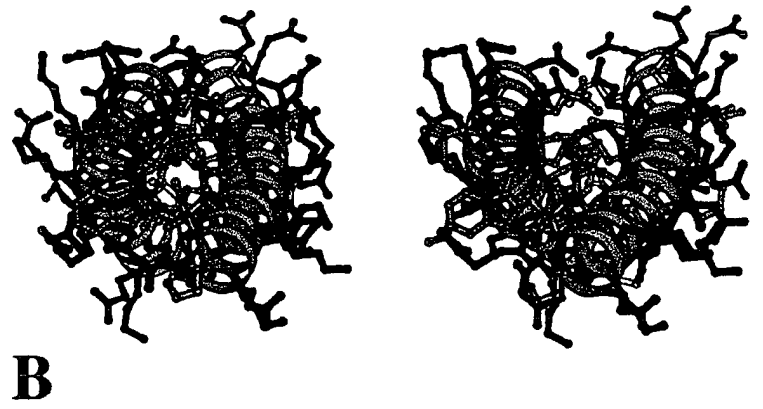

Figure 3B

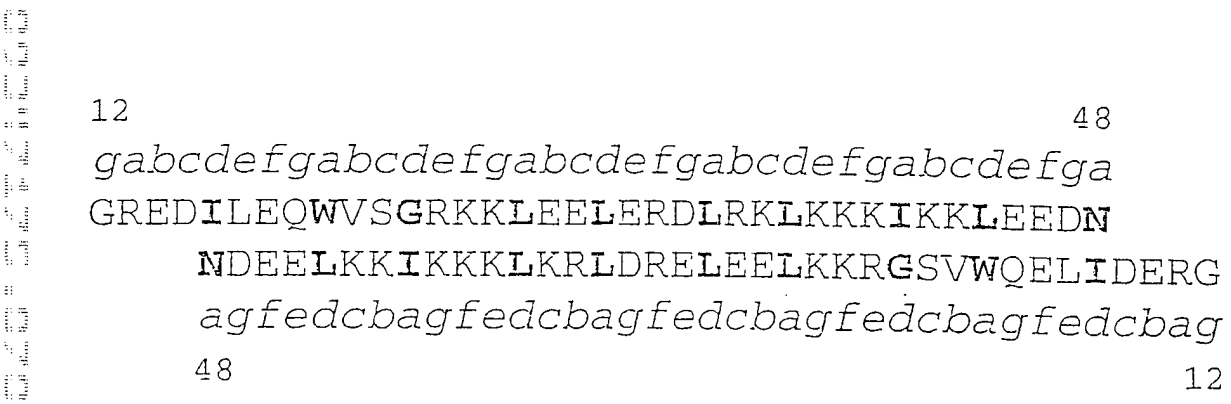
12
48
gabcdefgabcdefgabcdefgabcdefgabcdefga
GREDILEQWVSGRKKLEELERDLRKLKKKIKKLEEDN
NDEELKKIKKKLKRLDRELEELKKRGSVWQELIDERG
agfedcbagfedcbagfedcbagfedcbagfedcbag
48
12

Figure 3C

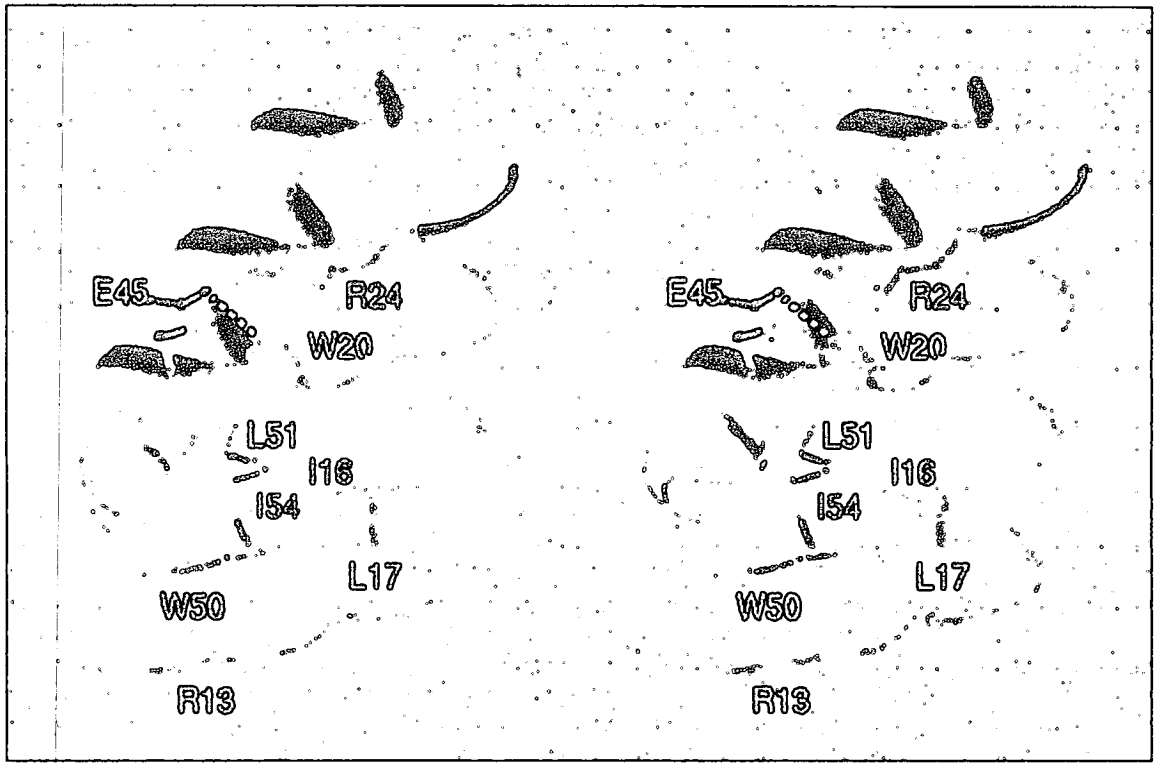
E45
R24
W20
L51
I16
I54
L17
W50
R13
E45
R24
W20
L51
I16
I54
L17
W50
R13

Figure 4

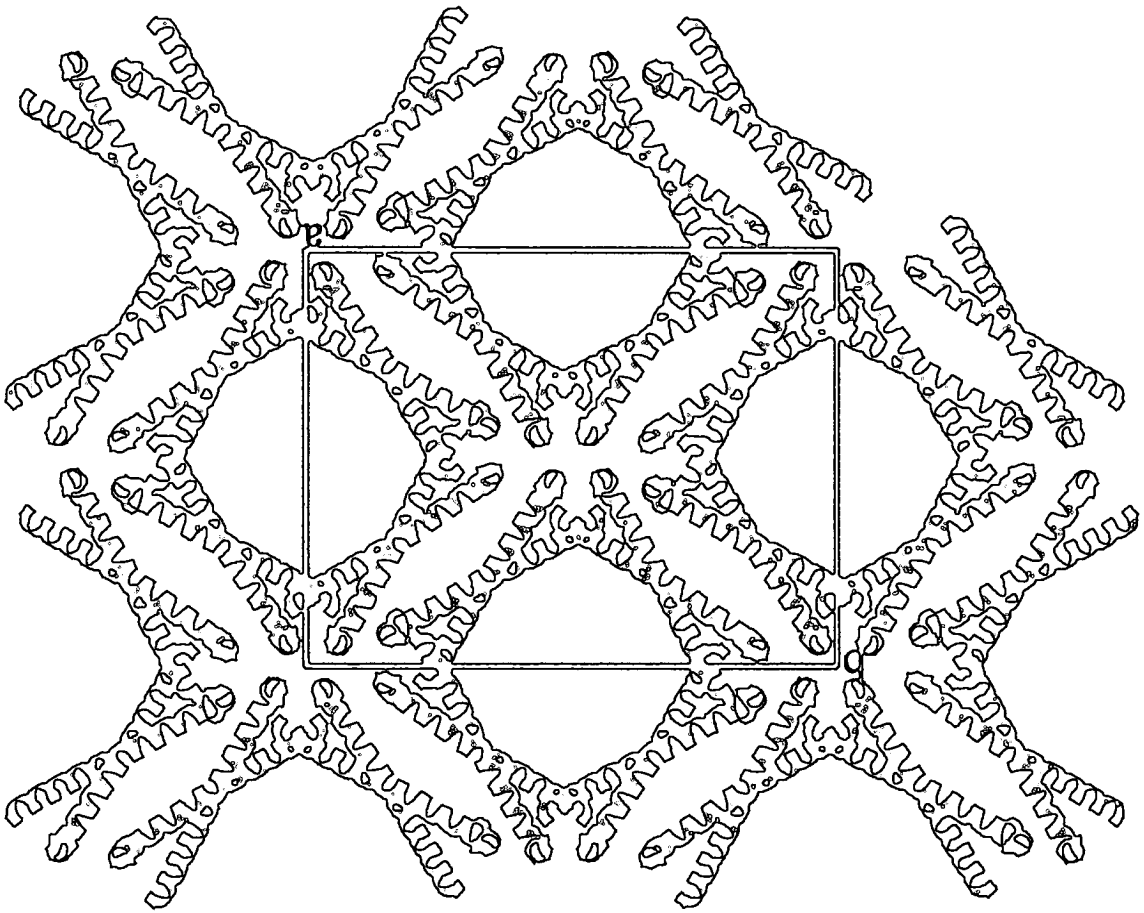

Figure 5

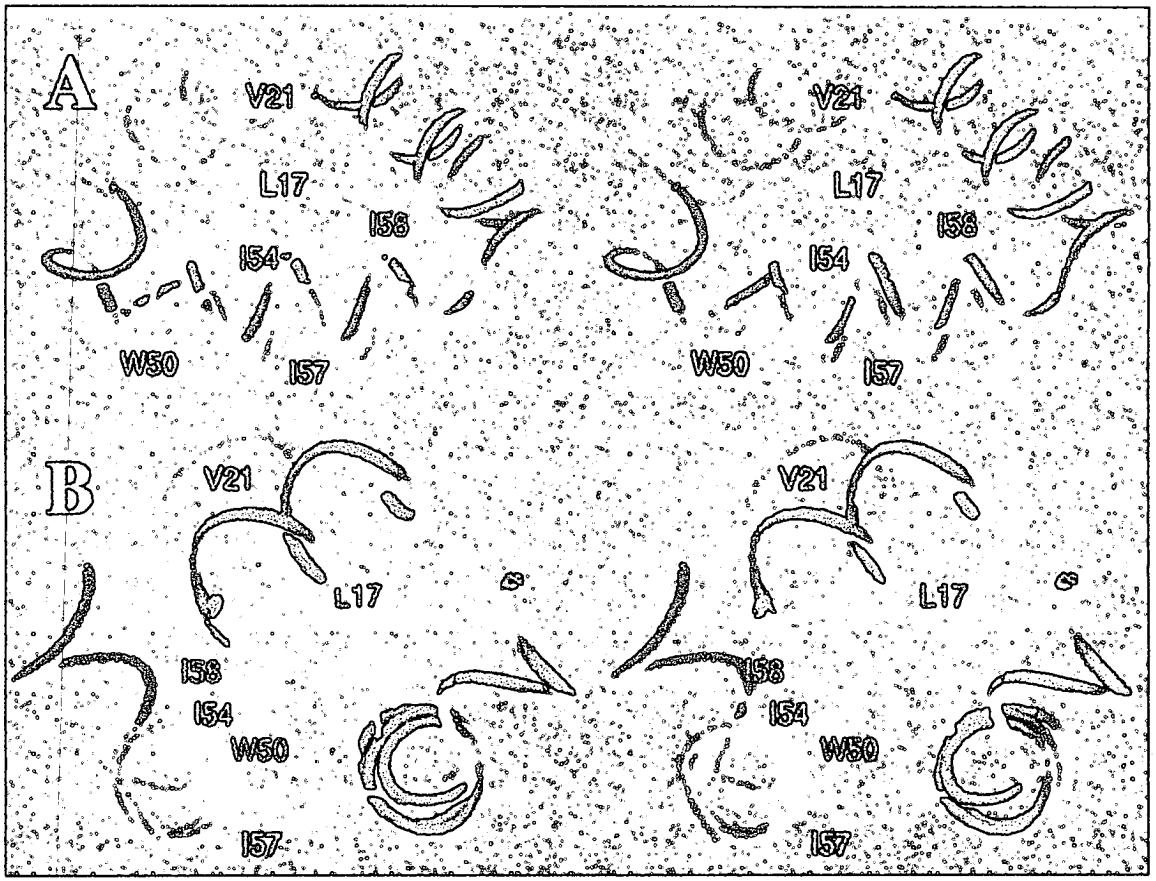

Figure 6A

Figure 6B

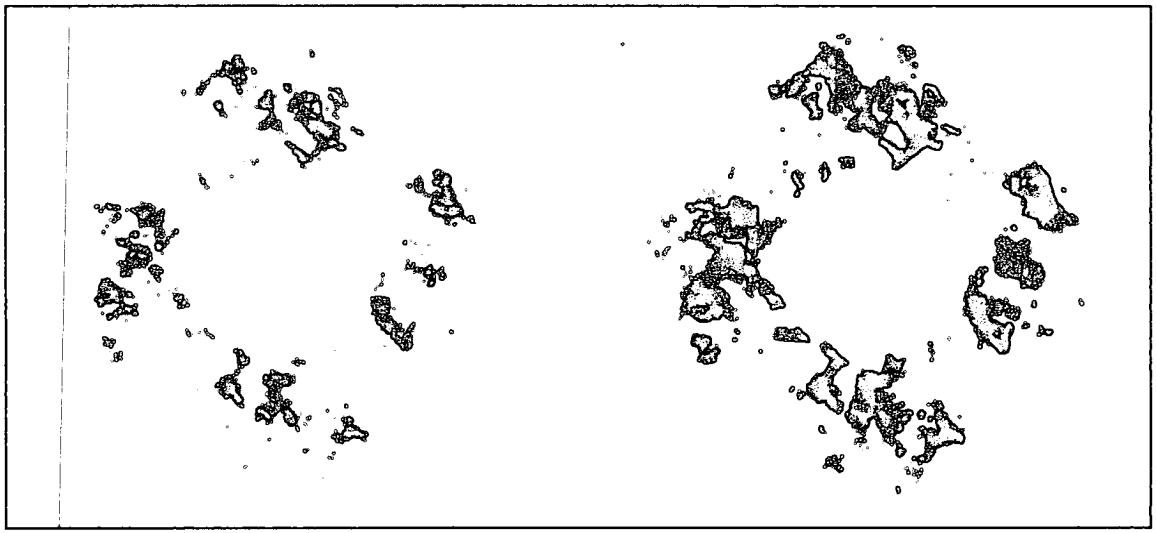

Figure 7A

Figure 7B

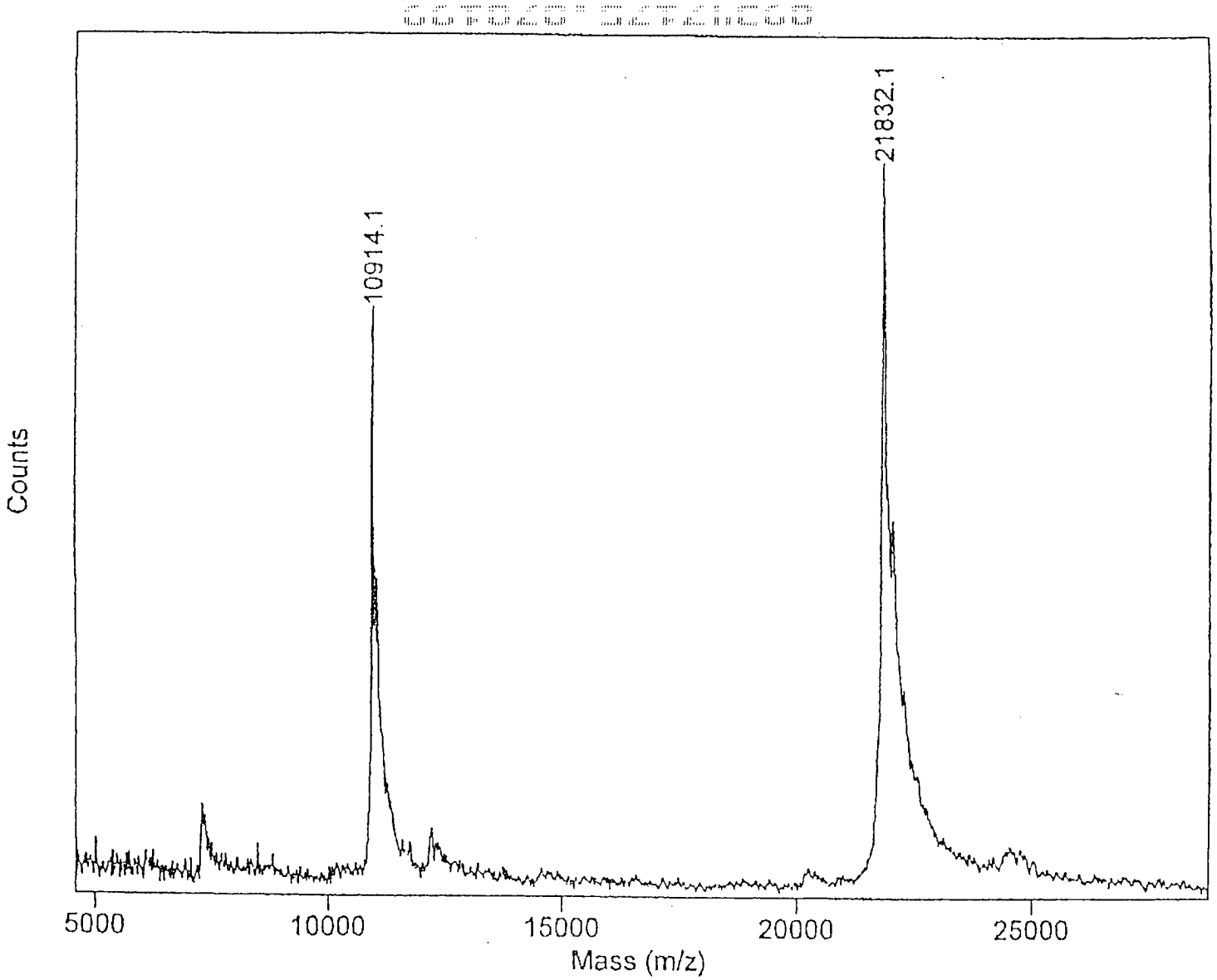

FIGURE 8A

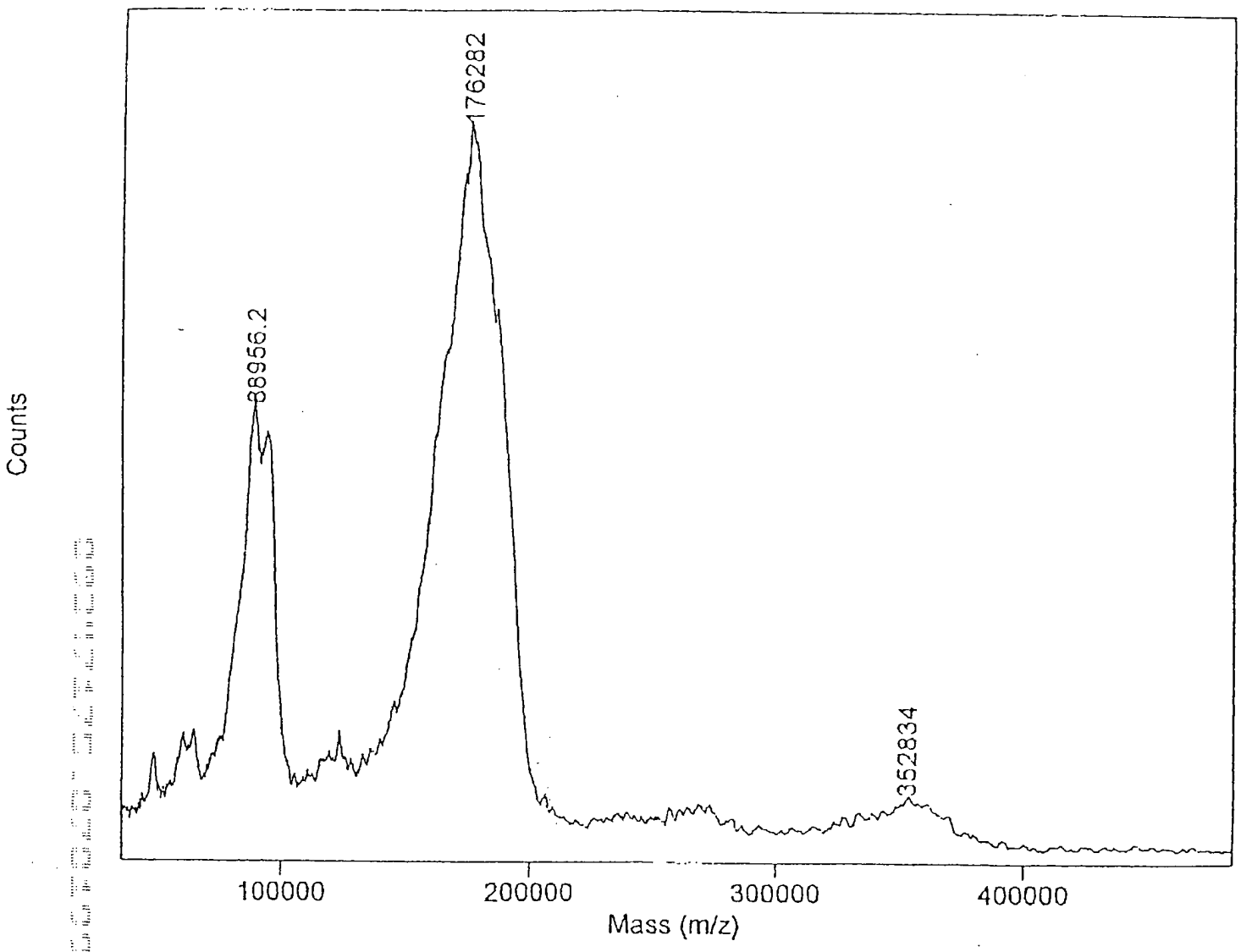
Counts
88956.2
176282
352834
100000
200000
300000
400000
Mass (m/z)

FIGURE 8B

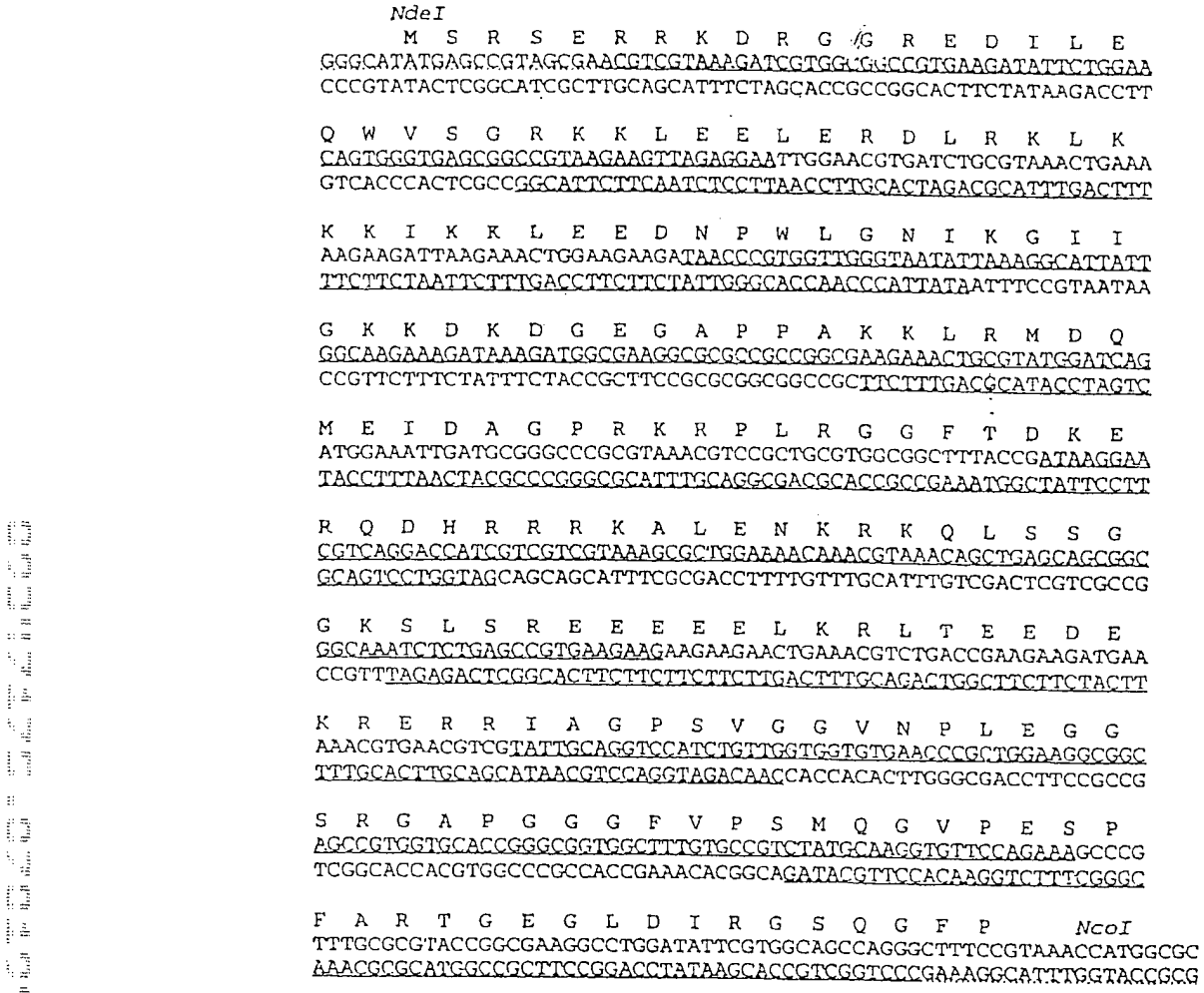
NdeI
M S R S E R R K D R G G R E D I L E
GGGCATATGAGCCGTAGCGAACGTCGTAAAGATCGTGGCGGCCGTGAAGATATTCTGGAA
CCCGTATACTCGGCATCGCTTGCAGCATTTCTAGCACCGCCGGCACTTCTATAAGACCTT
Q W V S G R K K L E E L E R D L R K L K
CAGTGGGTGAGCGGCCGTAAGAAGTTAGAGGAATTGGAACGTGATCTGCGTAAACTGAAA
GTCACCCACTCGCCGGCATTCTTCAATCTCCTTAACCTTGCACTAGACGCATTTGACTTT
K K I K K L E E D N P W L G N I K G I I
AAGAAGATTAAGAAACTGGAAGAAGATAACCCGTGGTTGGGTAATATTAAAGGCATTATT
TTCTTCTAATTCTTTGACCTTCTTCTATTGGGCACCAACCCATTATAATTTCCGTAATAA
G K K D K D G E G A P P A K K L R M D Q
GGCAAGAAAGATAAAGATGGCGAAGGCGCGCCGCCGGCGAAGAAACTGCGTATGGATCAG
CCGTTCTTTCTATTTCTACCGCTTCCGCGCGGCGGCCGCTTCTTTGACGCATACCTAGTC
M E I D A G P R K R P L R G G F T D K E
ATGGAAATTGATGCGGGCCCGCGTAAACGTCCGCTGCGTGGCGGCTTTACCGATAAGGAA
TACCTTTAACTACGCCCGGGCGCATTTGCAGGCGACGCACCGCCGAAATGGCTATTCCTT
R Q D H R R R K A L E N K R K Q L S S G
CGTCAGGACCATCGTCGTCGTAAAGCGCTGGAAAACAAACGTAAACAGCTGAGCAGCGGC
GCAGTCCTGGTAGCAGCAGCATTTCGCGACCTTTTGTTTGCATTTGTCGACTCGTCGCCG
G K S L S R E E E E E L K R L T E E D E
GGCAAATCTCTGAGCCGTGAAGAAGAAGAAGAACTGAAACGTCTGACCGAAGAAGATGAA
CCGTTTAGAGACTCGGCACTTCTTCTTCTTCTTGACTTTGCAGACTGGCTTCTTCTACTT
K R E R R I A G P S V G G V N P L E G G
AAACGTGAACGTCGTATTGCAGGTCCATCTGTTGGTGGTGTGAACCCGCTGGAAGGCGGC
TTTGCACTTGCAGCATAACGTCCAGGTAGACAACCACCACACTTGGGCGACCTTCCGCCG
S R G A P G G G F V P S M Q G V P E S P
AGCCGTGGTGCACCGGGCGGTGGCTTTGTGCCGTCTATGCAAGGTGTTCCAGAAAGCCCG
TCGGCACCACGTGGCCCCGCCACCGAAACACGGCAGATACGTTCCACAAGGTCTTTCGGGC
F A R T G E G L D I R G S Q G F P NcoI
TTTGCGCGTACCGGCGAAGGCCTGGATATTCGTGGCAGCCAGGGCTTTCCGTAAACCATGGCGC
AAACGCGCATGGCCGCTTCCGGACCTATAAGCACCGTCGGTCCCGAAAGGCATTTGGTACCGCG

Figure 9

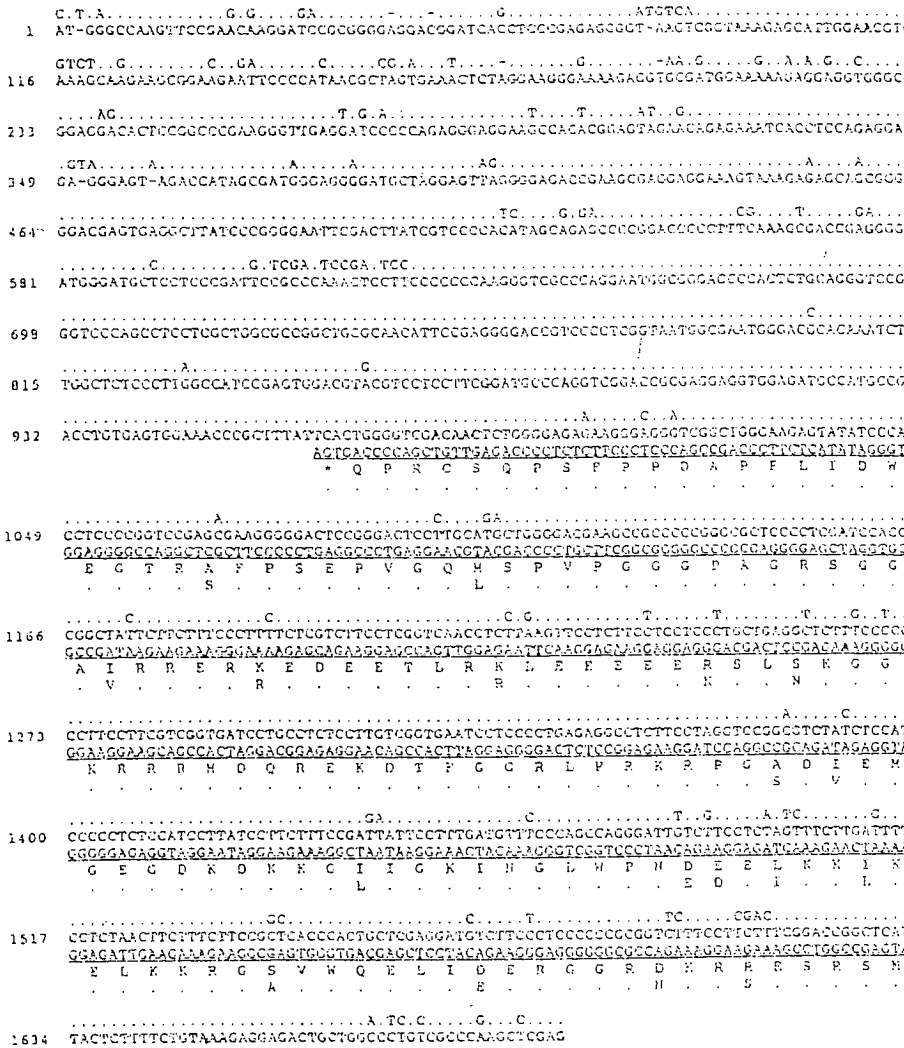

Figure 10

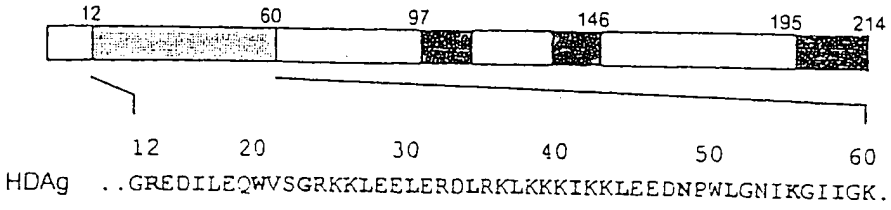

Figure 11A

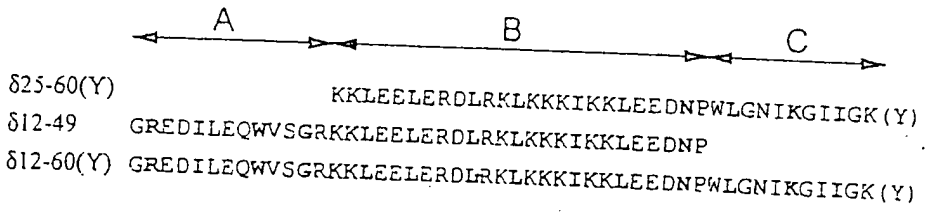

Figure 11B

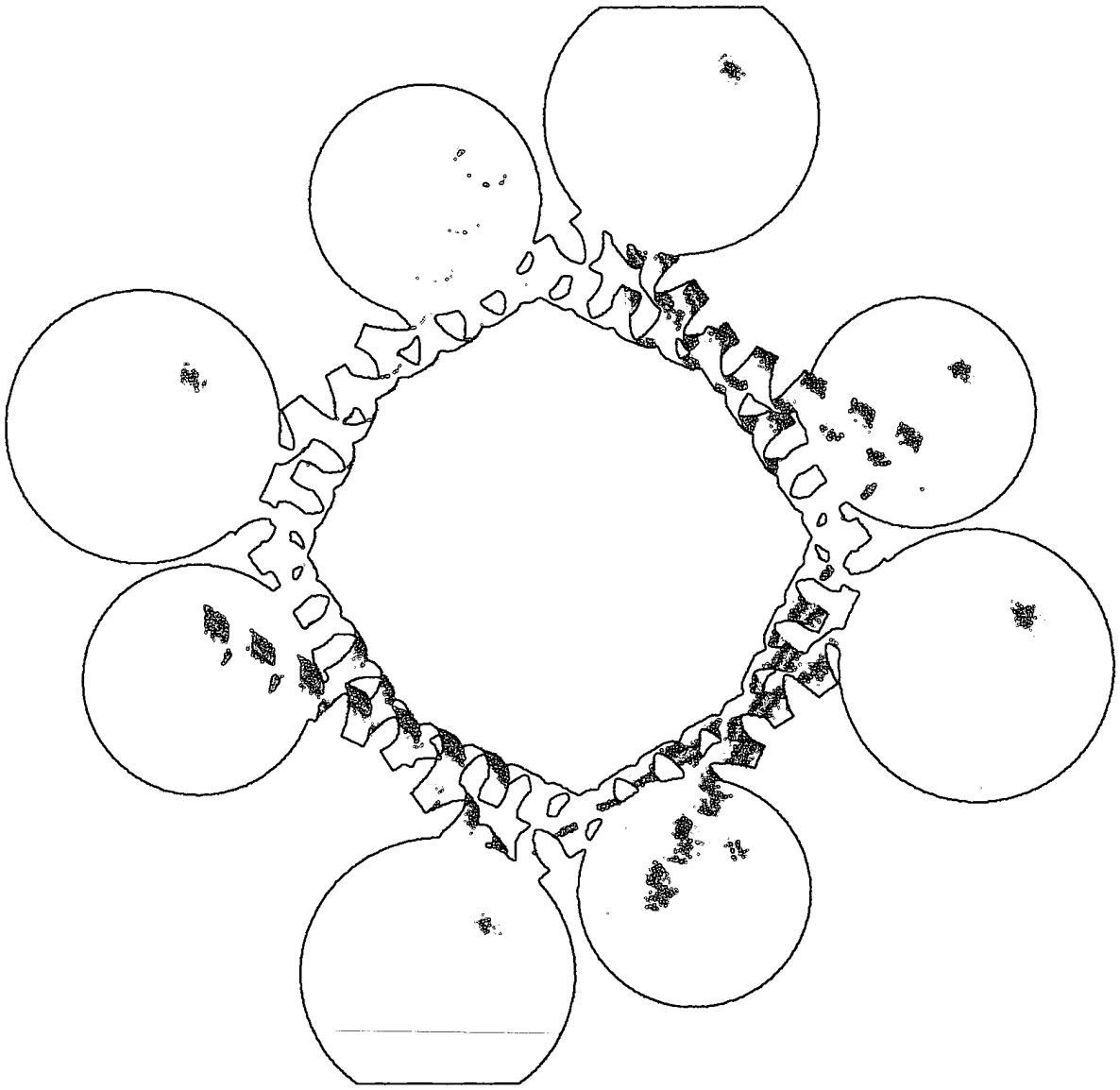

Figure 12A

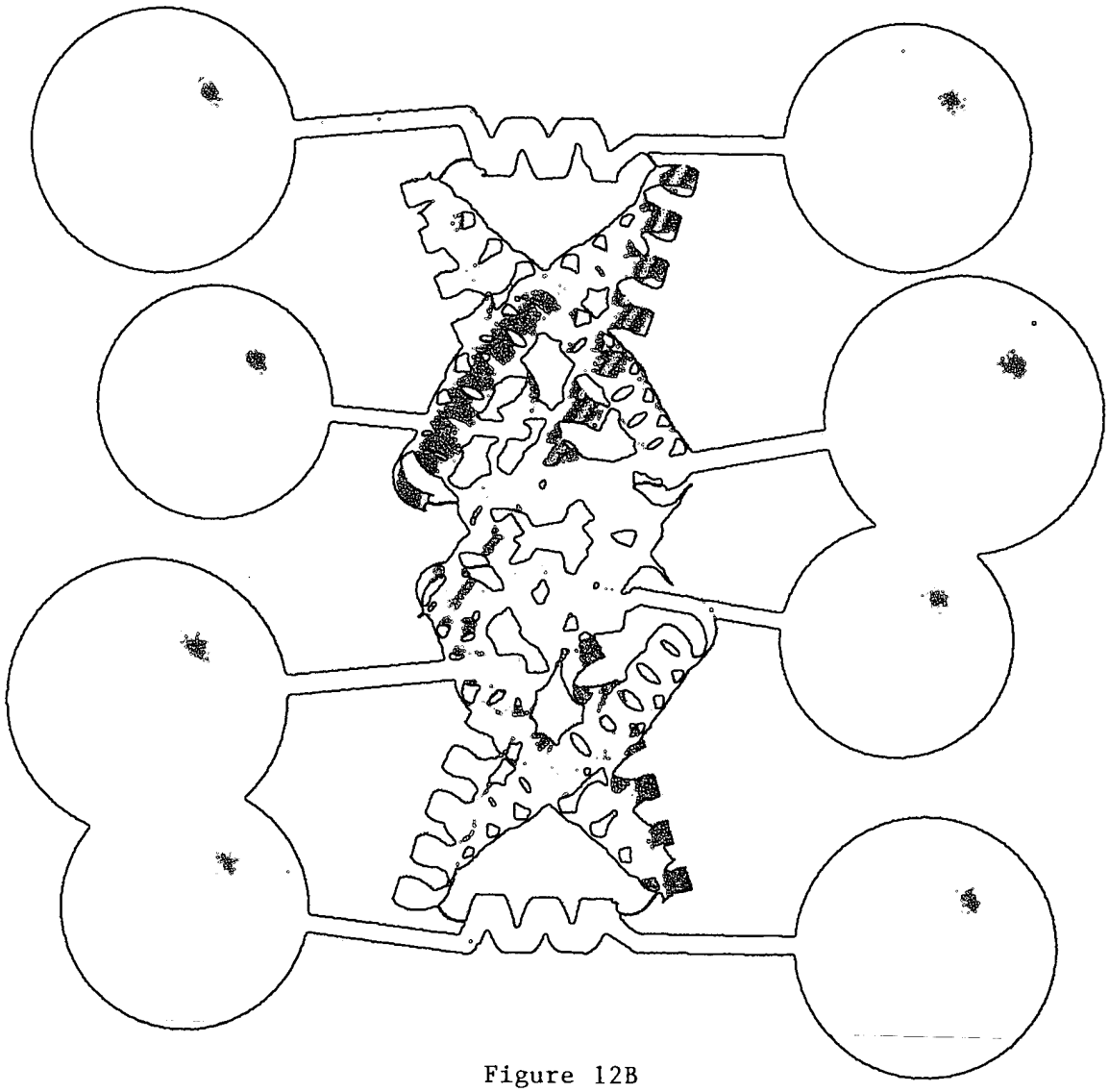

Figure 12B

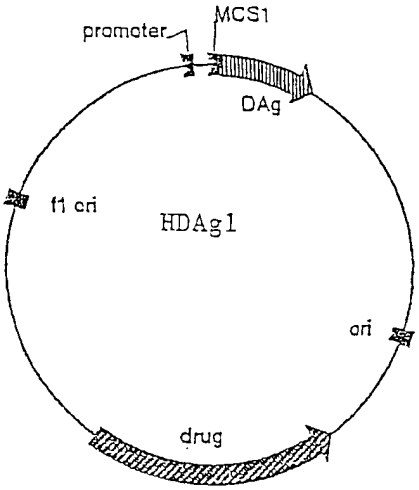

Figure 13a

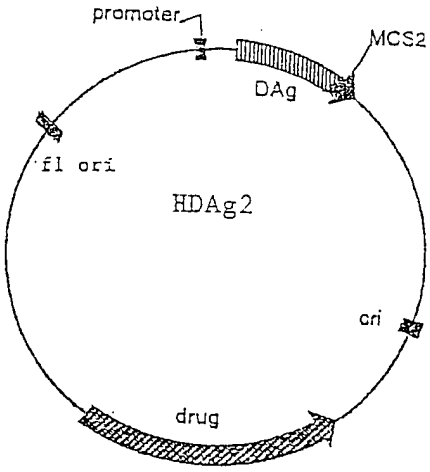

Figure 13b

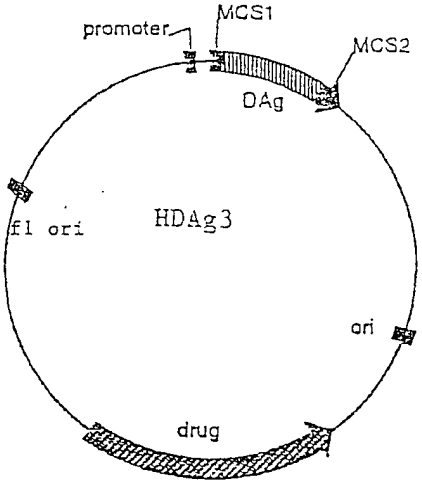

Figure 13c

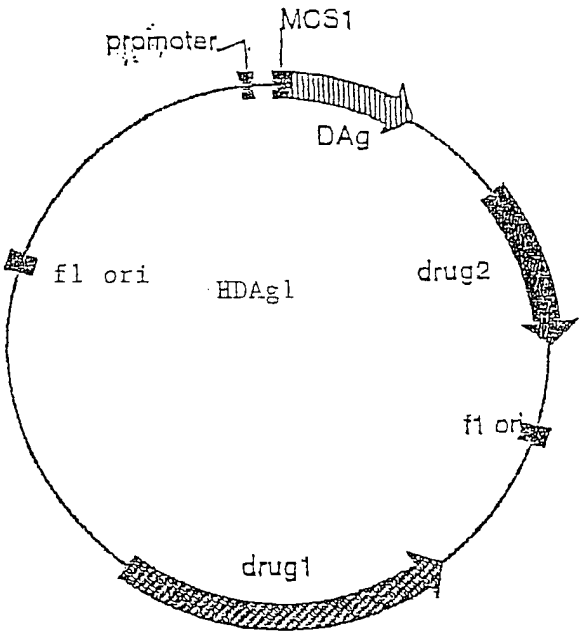

Figure 14

```
                    1                                                   50
sythetic ORF        ATGAGCCGta gCGAAcGtcG tAAAGAtCGt GGcGGccGtG AAGAtATtCT
wildtype ORF        ATGAGCCGgt cCGAAaGaaG gAAAGAcCGc GGgGGgaGgG AAGAcATcCT
          Identity  ATGAGCCG-- -CGAA-G--G -AAAGA-CG- GG-GG--G-G AAGA-AT-CT

                    51                                                  100
sythetic ORF        gGAaCAGTGG GTGAGCGGcc GtAAGAAGTT AGAGGAAtTg GAacGtGAtC
wildtype ORF        cGAgCAGTGG GTGAGCGGaa GaAAGAAGTT AGAGGAAcTc GAgaGaGAcC
          Identity  -GA-CAGTGG GTGAGCGG-- G-AAGAAGTT AGAGGAA-T- GA--G-GA-C

                    101                                                 155
sythetic ORF        TgCGtAAacT gAAaAAGAAg ATtAAGAAAC TgGAaGAAGA tAAcCCgTGG
wildtype ORF        TcCGgAAgtT aAAgAAGAAa ATcAAGAAAC TaGAgGAAGA cAAtCCcTGG
          Identity  T-CG-AA--T -AA-AAGAA- AT-AAGAAAC T-GA-GAAGA -AA-CC-TGG

                    151                                                 205
sythetic ORF        tTGGGtAAtA TtAAAGGcAT tATtGGcAAG AAaGATAAaG ATGGcGAaGG
wildtype ORF        cTGGGaAAcA TcAAAGGaAT aATcGGaAAG AAgGATAAgG ATGGaGAgGG
          Identity  -TGGG-AA-A T-AAAGG-AT -AT-GG-AAG AA-GATAA-G ATGG-GA-GG

                    201                                                 255
sythetic ORF        cGCgCCgCCG GCGAAGAAaC TgCGtATGGA tCAGATGGAa ATtGAtGCgG
wildtype ORF        gGCaCCcCCG GCGAAGAAgC TcCGgATGGA cCAGATGGAg ATaGAcGCcG
          Identity  -GC-CC-CCG GCGAAGAA-C T-CG-ATGGA -CAGATGGA- AT-GA-GC-G

                    251                                                 305
sythetic ORF        GcCCgcGtAA acGtCCgCTg cGtGGcGGcT TtACCGAtAA GGAacGtCAG
wildtype ORF        GaCCtaGgAA gaGgCCtCTc aGgGGaGGaT TcACCGAcAA GGAgaGgCAG
          Identity  G-CC--G-AA --G-CC-CT- -G-GG-GG-T T-ACCGA-AA GGA--G-CAG

                    301                                                 355
sythetic ORF        GAcCAtCGtC GtcGtAAaGC gCTgGAaAAC AAacGtAAaC AGCTgagcag
wildtype ORF        GAtCAcCGaC GaaGgAAgGC cCTcGAgAAC AAgaGgAAgC AGCTatcgtc
          Identity  GA-CA-CG-C G--G-AA-GC -CT-GA-AAC AA--G-AA-C AGCT------

                    351                                                 405
sythetic ORF        cGGcGGcAAa tctCTgAGCc GtGAaGAaGA AGAaGAACTg AAacGtcTGA
wildtype ORF        gGGgGGaAAg agcCTcAGCa GgGAgGAgGA AGAgGAACTt AAgaGgtTGA
          Identity  -GG-GG-AA- ---CT-AGC- G-GA-GA-GA AGA-GAACT- AA--G--TGA

                    401                                                 455
sythetic ORF        CCGAaGAAGA tGAaAAAcGt GAAcGtcGtA TtGCaGGtCC aTCtGTTGGt
wildtype ORF        CCGAgGAAGA cGAgAAAaGg GAAaGaaGaA TaGCcGGcCC gTCgGTTGGg
          Identity  CCGA-GAAGA -GA-AAA-G- GAA-G--G-A T-GC-GG-CC -TC-GTTGG-

                    451                                                 505
sythetic ORF        GGTGTGAACC CgCTgGAAGG cGGcagccGt GGtGCaCCgG GcGGtGGCTT
wildtype ORF        GGTGTGAACC CcCTcGAAGG tGGatcgaGg GGaGCgCCcG GgGGcGGCTT
          Identity  GGTGTGAACC C-CT-GAAGG -GG-----G- GG-GC-CC-G G-GG-GGCTT

                    501                                                 555
sythetic ORF        tGTgCCgtct ATGCAAGGtG TtCCaGAaag CCCgTTtGCg CGtACCGGcG
wildtype ORF        cGTcCCcagc ATGCAAGGaG TcCCgGAgtc CCCcTTcGCt CGgACCGGgG
          Identity  -GT-CC---- ATGCAAGG-G T-CC-GA--- CCC-TT-GC- CG-ACCGG-G

                    551                                                 603
sythetic ORF        AaGGcCTGGA tATtcGtGGc AGCCAGGGcT TtCCgTaaac cATggcgc
wildtype ORF        AgGGaCTGGA cATaaGgGGa AGCCAGGGaT TcCCaTggga tATactct
          Identity  A-GG-CTGGA -AT--G-GG- AGCCAGGG-T T-CC-T---- -AT-----
```

Figure 15

```
  1  GGGCATATGA GCCGTAGCGA ACGTCGTAAA GATCGTGGCG GCCGTGAAGA
 51  TATTCTGGAA CAGTGGGTGA GCGGCCGTAA GAAGTTAGAG GAATTGGAAC
101  GTGATCTGCG TAAACTGAAA AAGAAGATTA AGAAACTGGA AGAAGATAAC
151  CCGTGGTTGG GTAATATTAA AGGCATTATT GGCAAGAAAG ATAAAGATGG
201  CGAAGGCGCG CCGCCGGCGA AGAAACTGCG TATGGATCAG ATGGAAATTG
251  ATGCGGGCCC GCGTAAACGT CCGCTGCGTG GCGGCTTTAC CGATAAGGAA
301  CGTCAGGACC ATCGTCGTCG TAAAGCGCTG GAAAACAAAC GTAAACAGCT
351  GAGCAGCGGC GGCAAATCTC TGAGCCGTGA AGAAGAAGAA GAACTGAAAC
401  GTCTGACCGA AGAAGATGAA AAACGTGAAC GTCGTATTGC AGGTCCATCT
451  GTTGGTGGTG TGAACCCGCT GGAAGGCGGC AGCCGTGGTG CACCGGGCGG
501  TGGCTTTGTG CCGTCTATGC AAGGTGTTCC AGAAAGCCCG TTTGCGCGTA
551  CCGGCGAAGG CCTGGATATT CGTGGCAGCC AGGGCTTTCC GTAAACCATG
601  GCGC
```

Figure 16

```
                  1                                                   48
wildtype HDAg-S   MSRSERRK DRGGREDILE QWVSGRKKLE ELERDLRKLK KKIKKLEEDN
pR5DV5 plasmid    MSRSERRK DRGGREDILE QWVSGRKKLE ELERDLRKLK KKIKKLEEDN
       Identity   MSRSERRK DRGGREDILE QWVSGRKKLE ELERDLRKLK KKIKKLEEDN

                  49                                                  98
wildtype HDAg-S   PWLGNIKGII GKKDKDGEGA PPAKKLRMDQ MEIDAGPRKR PLRGGFTDKE
pR5DV5 plasmid    PWLGNIKGII GKKDKDGEGA PPAKKLRMDQ MEIDAGPRKR PLRGGFTDKE
       Identity   PWLGNIKGII GKKDKDGEGA PPAKKLRMDQ MEIDAGPRKR PLRGGFTDKE

                  99                                                 148
wildtype HDAg-S   RQDHRRRKAL ENKRKQLSSG GKSLSREEEE ELKRLTEEDE KRERRIAGPS
pR5DV5 plasmid    RQDHRRRKAL ENKRKQLSSG GKSLSREEEE ELKRLTEEDE KRERRIAGPS
       Identity   RQDHRRRKAL ENKRKQLSSG GKSLSREEEE ELKRLTEEDE KRERRIAGPS

                  149                                                195
wildtype HDAg-S   VGGVNPLEGG SRGAPGGGFV PSMQGVPESP FARTGEGLDI RGSQGFP
pR5DV5 plasmid    VGGVNPLEGG SRGAPGGGFV PSMQGVPESP FARTGEGLDI RGSQGFP
       Identity   VGGVNPLEGG SRGAPGGGFV PSMQGVPESP FARTGEGLDI RGSQGFP
```

Figure 17

primer1
GGGCATATGAGCCGTAGCGAACGTCGTAAAGATCGTGGCGGCCGTGAAGATA
TTCTGGAACAGTGGGTGAGCGGCCGTAAGAAGTTAGAGGAA

primer2
ATATTACCCAACCACGGGTTATCTTCTTCCAGTTTCTTAATCTTCTTTTT
CAGTTTACGCAGATCACGTTCCAATTCCTCTAACTTCTTACGGCC

primer3
TAACCCGTGGTTGGGTAATATTAAAGGCATTATTGGCAAGAAAGATAAAG
ATGGCGAAGGCGCGCCGCCGGCGAAGAAACTGCGTATGGATCAG

primer4
GATGGTCCTGACGTTCCTTATCGGTAAAGCCGCCACGCAGCGGACGTTTA
CGCGGGCCCGCATCAATTTCCATCTGATCCATACGCAGTTTCTT

primer5
ATAAGGAACGTCAGGACCATCGTCGTCGTAAAGCGCTGGAAAACAAACGT
AAACAGCTGAGCAGCGGCGGCAAATCTCTGAGCCGTGAAGAAG

primer6
CAACAGATGGACCTGCAATACGACGTTCACGTTTTTCATCTTCTTCGGTC
AGACGTTTCAGTTCTTCTTCTTCTTCACGGCTCAGAGAT

primer7
TATTGCAGGTCCATCTGTTGGTGGTGTGAACCCGCTGGAAGGCGGCAGCC
GTGGCGCGCCGGGCGGCGGCTTTGTGCCGTCTATGCAAGGTGTTCCAGAA
A

primer8
GCGCCATGGTTTACGGAAAGCCCTGGCTGCCACGAATATCCAGGCCTTCG
CCGGTACGCGCAAACGGGCTTTCTGGAACACCTTGCATAG

primer9
GGGCATATGAGCCGTAGCGA

primer10
GCGCCATGGTTTACGGAAAG

Figure 18